Claims

1. Method for pre-processing speech, in particular in a method for recognizing speech, comprising the steps of:
- receiving a speech signal (S),
- separating a spectrum (F) of said speech signal (S) into a given number (N) of predetermined frequency sub-bands (F_1, ..., F_N),
- analyzing said speech signal (S) within each of said frequency sub-bands (F_1, ..., F_N),
- thereby generating respective band-dependent acoustic feature data (O_1, ..., O_N) for each of said respective frequency sub-bands (F_1, ..., F_N), which band-dependent acoustic feature data (O_1, ..., O_N) are at least in part representative for said speech signal (S) with respect to a respective frequency sub-band (F_1, ..., F_N),
- deriving band-dependent likelihoods (b_1, ..., b_N) for occurrences of speech elements (P1, ..., Pm) or of sequences thereof within said speech signal (S) based on said band-dependent acoustic feature data (O_1, ..., O_N) and/or a derivative thereof,
- analyzing said speech signal (S) within said entire spectrum (F),
- thereby generating full-band acoustic feature data (FBE-F; FFBE; FBE-F-SSUB; $O_{F,SSUB}$), which are at least in part representative for said speech signal (S) with respect to said entire spectrum (F),
- deriving a full-band likelihood (B_{FF}; B_{SSUB}) for occurrences of speech elements (P1, ..., Pm) or of sequences thereof within said speech signal (S) based on said full-band acoustic feature data (FBE-F; FFBE; FBE-F-SSUB; $O_{F,SSUB}$) and/or a derivative thereof,
- deriving an overall likelihood (B) for occurrences of speech elements (P1, ..., Pm) or of sequences thereof within said speech signal (S) based on said band-dependent likelihoods (b_1, ..., b_N) and said full-band likelihood (B_{FF}; B_{SSUB}).

2. The method according to claim 1,
characterized in that
when deriving said overall likelihood (B) said band-dependent likelihoods (b_1, ..., b_N) are combined to a union model likelihood ($B_{U.MFCC}$) by determining the number of uncorrupted frequency sub-bands of said frequency sub-bands (F_1, ..., F_N), and adding all possible combinations of products of different

band-dependent likelihoods (b_1, ..., b_N) corresponding to respective frequency sub-bands.

3. The method according to any one of the preceding claims,
characterized in that
said band-dependent acoustic feature data (O_1, ..., O_N) comprise respective band-dependent mel-frequency cepstral coefficient features, which are based on mel-frequency cepstral coefficients and/or a derivative thereof derived from respective frequency sub-bands (F_1, ..., F_N).

4. The method according to any one of the preceding claims,
characterized in that
a predetermined broadband noise robustness technique is applied prior to deriving said full-band likelihood term (B_{FF}; B_{SSUB}).

5. The method according to claim 4,
characterized in that
said broadband noise robustness technique is based on a frequency-filtering technique.

6. The method according to claim 4,
characterized in that
said broadband noise robustness technique is based on a method of spectral-subtraction.

7. The method according to any one of the preceding claims,
characterized in that
said full-band acoustic feature data (FBE-F; FFBE; FBE-F-SSUB; $O_{F,SSUB}$) comprise filter bank energy features (FBE-F), which are based on filter bank energies derived from said entire spectrum (F).

8. The method according to any one of the preceding claims,
characterized in that
said full-band acoustic feature data (FBE-F; FFBE; FBE-F-SSUB; $O_{F,SSUB}$) comprise filtered filter bank energy features (FFBE), which are based on filtered filter bank energies derived from said entire spectrum (F).

9. The method according to any one of the preceding claims,

characterized in that
said full-band acoustic feature data (FBE-F; FFBE; FBE-F-SSUB; $O_{F.SSUB}$) comprise full-band mel-frequency cepstral coefficient features, which are based on mel-frequency cepstral coefficients and/or a derivative thereof derived from said entire spectrum (F).

10. The method according to any one of the preceding claims,
characterized in that
said full-band acoustic feature data (FBE-F; FFBE; FBE-F-SSUB; $O_{F.SSUB}$) and/or said band-dependent acoustic feature data (O_1, ..., O_N) comprise PLP-linear prediction filter features, which are based on PLP-linear prediction filter coefficients.

11. The method according to any one of the preceding claims,
characterized in that
said full-band acoustic feature data (FBE; FFBE; FBE-F-SSUB; $O_{F.SSUB}$) comprise spectrally-changed full-band mel-frequency cepstral coefficient features ($O_{F.SSUB}$), which are generated by applying a method of spectral subtraction to said full-band mel-frequency cepstral coefficient features (O_F).

12. The method according to any one of the preceding claims,
characterized in that
said band-dependent likelihoods (b_1, ..., b_N) and said likelihood term (B_{FF}; B_{SSUB}; $B_{U.FF}$) are determined using a probability estimator.

13. The method according to any one of the preceding claims,
characterized in that
said filtered filter bank energies (FFBE) are derived from said filter bank energies (FBE) by subtracting (f(i) = f(i+1) – f(i-1)) a first filter bank energy (FBE_{i-1}) from a second filter bank energy (FBE_{i+1}), wherein said first filter bank energy (FBE_{i-1}) corresponds to a first discrete frequency and said second filter bank energy (FBE_{i+1}) corresponds to a second discrete frequency, lying two discrete frequency steps after said first filter bank energy (FBE_{i-1}).

14. Speech pre-processing system, in particular integrated into a speech processing system, which is capable of performing or realizing a method for pre-

processing speech according to any one of the preceding claims 1 to 13 and/or the steps thereof.

15. Computer program product,
comprising computer program means adapted to perform and/or to realize the method of pre-processing speech according to any one of the claims 1 to 13 and/or the steps thereof, when it is executed on a computer, a digital signal processing means, and/or the like.

16. Computer readable storage medium,
comprising a computer program product according to claim 15.